# 6 Using Statistical Models to Study Temporal Dynamics of Animal–Landscape Relations

KEVIN J. GUTZWILLER AND SAMUEL K. RIFFELL

**Abstract.** Temporal variation in animal responses to landscape conditions may affect animal distributions, population and community structure, and resource use. Measuring such variation and studying its influence is essential for developing a realistic understanding of animal–landscape relations. Several statistical modeling approaches are appropriate for explicitly incorporating time into analyses of animal–landscape relations, but landscape ecologists have not commonly used them. Analytical assessment of temporal variation in animal–landscape relations may involve independent or dependent data. In the case of independent data, interaction effects involving time and landscape metrics can be estimated using cross-product terms. This approach permits direct comparison of animal–landscape regression curves across levels of time, enabling one to infer explicitly how relations vary temporally. With dependent (repeated measures) data, analytical assessment of temporal variation in animal–landscape relations may involve one (time), two (space, time), or three (two space, one time) dimensions. Independent-error methods to test for differences among means or regression curves are not valid in these situations. When data are recorded at equal time intervals, covariance structures that reflect correlations among observations that decrease with time, such as the autoregressive structure, can be used. When data are recorded at unequal time intervals, appropriate covariance structures include the power law, Gaussian, and spherical structures. A mixed-model approach can be used to draw inferences about interactions involving time and landscape metrics when one-, two-, and three-dimensional repeated measures are involved. In summary, several methods accessible to those with moderate training in statistics can be used to incorporate time into studies of animal–landscape relations. Land-use planning and biological conservation will benefit greatly from a better understanding of the temporal aspects of such relations.

## 6.1. Introduction

The study of animal–landscape relations has mushroomed in recent years as ecologists and conservation biologists have begun to appreciate more fully the potential effects of the surrounding landscape on animals at local sites. This landscape perspective is improving understanding of the broad-scale factors that influence species richness, abundance, population dynamics, and habitat use. Because these advances are leading to better representations of reality, their application may improve conservation effectiveness (Gutzwiller, 2002).

Landscapes and animal populations are dynamic. Disturbances such as plant diseases, storms, fire, floods, commercial and residential development, agricultural development, road construction, and silviculture alter the structure and composition of landscapes at specific extents. The time frame for such changes can range from days to decades. Animal populations can fluctuate seasonally and yearly with changes in food availability and quality, the quality and quantity of habitat, survival and reproduction driven by unpredictable weather, and life-history factors such as nomadic behavior and events occurring on distant wintering ranges.

Together, these sources of variation can lead to temporal variation in animal–landscape relations. For instance, habitats can be population sources in some years, but sinks in others (McCoy et al., 1999). Substantial year-to-year variation occurs in Great Lakes water levels, and coastal wetlands can be inundated one year but not the next, resulting in different animal–landscape relations in successive years (Riffell et al., 2003). Density dependence has been implicated as a source of temporal variation in animal–habitat relations (O'Connor, 1986). Even during periods when broad-scale habitat conditions are stable, considerable interannual variation in bird–landscape relations can occur (Gutzwiller and Barrow, 2001, 2002). Many examples exist for seasonal differences in habitat use, which may reflect niche shifts or niche extensions (Shochat and Tsurim, 2004 and references therein).

If we hope to understand and predict animal–landscape relations, our analytical approaches must incorporate temporal variation in explicit and robust ways. Temporal variation in animal-habitat relations—including relations at broad spatial scales—is poorly understood and in need of immediate study (Morrison, 2002). Compared to studies that do not address temporal factors, studies that explicitly include time are likely to yield information about animal–landscape relations that is more realistic and hence more useful to land-use planners and conservation biologists.

Several statistical approaches are appropriate for explicitly incorporating time into analyses of animal–landscape relations, but landscape ecologists have not commonly used them. The statistical modeling methods we discuss are well-established, and we anticipate their use will increase substantially once landscape ecologists become familiar with their utility and ease of application. The primary goal of this chapter is to increase understanding and application of these techniques so that temporal influences are more frequently incorporated into studies of animal–landscape relations.

## 6.2. Objectives

To accomplish this goal, we define uncommon statistical terms used in this chapter, explain when techniques for analyzing independent data and dependent data should be applied, and demonstrate statistical-modeling approaches for studying temporal variation in animal–landscape relations. Researchers with modest statistical training—knowledge of basic analysis of variance and regression, for example—can implement the modeling techniques we consider. We use SAS software (SAS Institute, 2002) in our examples because it is widely available, it can be used to model many types of response variables of interest to landscape ecologists, it enables one to model a large array of temporal and spatial covariance structures, and it is the platform with which we are most familiar.

We explain how to use SAS for five statistical modeling approaches that explicitly incorporate time: (1) time-related interaction terms in regression models using independent observations; (2) mixed models for temporally dependent observations that are equally spaced in time; (3) mixed models for temporally dependent observations that are not equally spaced in time; (4) mixed models for temporally and spatially dependent observations; and (5) mixed models for data that exhibit dependence in two spatial dimensions and one temporal dimension. To improve understanding of the techniques discussed in this chapter, we provide simplified definitions of statistical terms (Box 6.1).

**Box 6.1. Definitions of statistical terms.**

*Autocorrelation*—correlation between temporally or spatially successive observations of a variable in a data set.

*Covariance structures*—different patterns of correlation among observations from the same or different sampling units.

*Cross-product*—the result of multiplying the values of two explanatory variables together for a particular sampling unit. Cross-products for an entire sample can be used as the data for estimating the coefficient for an interaction variable in regression models.

*Fixed effect*—an effect whose levels in an analysis represent all possible levels, or at least all of the levels about which inference is to be made.

*Full model*—the most complex mean model under consideration, containing all fixed effects of interest.

*Maximum likelihood (ML)*—a method of estimating parameter values based on maximizing the likelihood function.

*Mixed model*—a model containing both fixed and random effects.

*Random effect*—an effect whose levels in an analysis represent a random subset of the possible levels.

*Repeated measures*—multiple observations obtained from the same sampling unit (e.g., plot, animal, station) in sequence over time. This term also is used to describe types of analyses designed to accommodate such data (e.g., repeated measures analysis of variance).

*Restricted maximum likelihood (REML)*—a method of parameter estimation restricted to maximizing the likelihood function over the random effects portion of a model.

## 6.3. Assessing Temporal Variation in Animal–Landscape Relations Using Independent Observations

### *6.3.1. Independent Data in Landscape Studies*

In a number of research situations, animal metrics (e.g., species richness, abundance, habitat use) may be measured in different landscapes over time. Time frames may include a single season, multiple seasons, or different years. It is not always feasible to gather synchronous observations in many landscapes, or multiple observations through time in each of many landscapes. Remote locations, and constraints on personnel or time available for research, for example, can prevent simultaneous or near-simultaneous surveys of all landscapes. The result can be one measure of the response variable for each of many separate landscapes but across a span of time (e.g., Pearson, 1993; McGarigal and McComb, 1995; Naugle et al., 1999).

For instance, we may need to study mammal–landscape relations during a breeding season based on asynchronous surveys in different landscapes. But mammals might occupy landscapes differently as the season progresses because of the phenology of plants, changes in temperature, or changes in other endogenous or exogenous factors to which mammals respond. Under these circumstances, assessment of the relation between mammals and landscape features would be misleading if time was influential but was left out of the analysis; i.e., if mammal–landscape relations varied with the time of the season, it would be essential to explicitly incorporate time into the modeling process.

As another example, consider a scenario in which snake density was sampled in numerous landscapes during a two-year study. Not enough funding was available to survey any landscape more than once. Instead of obtaining multiple observations over time in the same landscapes, the investigators decided to allocate their resources in a way that would provide information about a larger number of landscapes. This decision was motivated in part by available resources but also by the desire to include a wide range of landscape conditions in the analyses so that any resulting model would have greater potential for robust prediction in the study region. Accordingly, snake density was measured for half of the landscapes during the first year and for the other half during the second year. The landscapes were far enough apart that the estimates of snake density for the different landscapes were independent. Thus, the researchers had a set of independent observations with the potential to exhibit interaction effects between time and landscape features.

### *6.3.2. Interaction Effects*

When data for a response variable are collected at independent locations over time, temporal variation in animal–landscape relations can be studied by analyzing whether there are significant time-related interaction effects involving landscape variables. Returning to our snake research scenario, grassland cover was expected to be a key determinant of snake density, so the researchers used a geographic information system and digital land-cover data to measure percent grassland cover for each landscape. The study objective was to assess the relation between snake density and percent cover of grassland, but the relation between snake density and grassland cover may not have been the same during both years.

Specifically, the change in snake density per unit change in grassland cover (regression slope) may have differed between years. When the effect of an explanatory variable (e.g., grassland cover) on the response variable (e.g., snake density) varies with the level of another explanatory variable (e.g., year), an interaction effect (involving the two explanatory variables) exists on the response variable. Note that a grassland cover × year interaction effect would differ from a significant main effect for grassland cover (in which there would simply be a relation between snake density and grassland cover) and from a significant main effect for year (in which there would simply be a between-year difference in snake density).

With independent data, landscape ecologists can employ interaction terms in standard least-squares and logistic regression models to test whether there is a significant difference in animal–landscape relations over time. A convenient way to test for such dynamics is to calculate the cross-products (Neter et al., 1989) of a landscape and time metric. The cross-products are the data used in the analysis to test for an interaction effect. Any combination of discrete or continuous variables can be used to form the cross-product variable, and the regression coefficient associated with the cross-product variable represents the interaction term in the model. Multiple interaction terms can be examined in the same regression model. Examination of interaction effects enables one to infer how animal–landscape relations vary over time.

### *6.3.3. Example of SAS Code and Results*

Continuing with our snake example, the data for the interaction effect is the product of grassland cover multiplied by an indicator for year; the first year is represented in the data set with a 1 and the second year is represented with a 0. Coding of indicator variables is a common technique in regression (Neter et al., 1989). The data used in this example are available from the authors.

Using `lndscape` to represent landscape, `msnkden` to represent mean snake density, `grasscov` to represent grassland cover in the landscape, `year` to represent the year when mean snake density was measured, and `grssxyr` to represent the cross-products for the grassland cover × year interaction, SAS code for a standard least-squares regression to test for the interaction would look like the following:

```
data snake;
input lndscape msnkden grasscov years;
cards;
1      0.25  10  1
2      0.20  15  1
. . .
. . .
39     0.42  51  0
40     0.35  39  0
run;
grssxyr = grasscov*year;/*calculating the cross-product*/
proc reg;
model msnkden = grasscov year grssxyr;
run;
```

To determine whether there is a significant interaction effect, we examine the table of parameter (regression coefficient) estimates in the output:

Parameter estimates

| Variable | DF | Parameter estimate | Standard error | *t* | Pr > \|*t*\| |
|---|---|---|---|---|---|
| intercept | 1 | 0.3272 | 0.0275 | 11.92 | <0.0001 |
| grasscov | 1 | 0.0014 | 0.0005 | 2.70 | 0.0106 |
| year | 1 | −0.1451 | 0.0376 | −3.86 | 0.0004 |
| grssxyr | 1 | 0.0039 | 0.0007 | 5.88 | <0.0001 |

The parameter estimate for the interaction term (`grssxyr`) is significantly different from zero, implying that the relation between mean snake density and grassland cover varies with year. We can visualize this result by plotting the relation between mean snake density and grassland cover for each year separately on the same graph (Fig. 6.1). When the regression lines in this type of graph are not parallel (slopes are not equal), there is evidence of an interaction (Neter et al., 1989). In our example, mean snake density increased with grassland cover, but it did so at a higher rate in year 1 compared to year 2. Thus, the animal–landscape relation exhibited temporal flux.

This approach to assessing interaction effects also can be applied in a general linear model context (using SAS's Proc GLM) and in a logistic regression setting (using SAS's Proc Logistic) (SAS Institute, 2002). In Proc GLM and Proc Logistic, an assignment statement to define the interaction is not needed before the model statement; the interaction term is specified in the model statement itself.

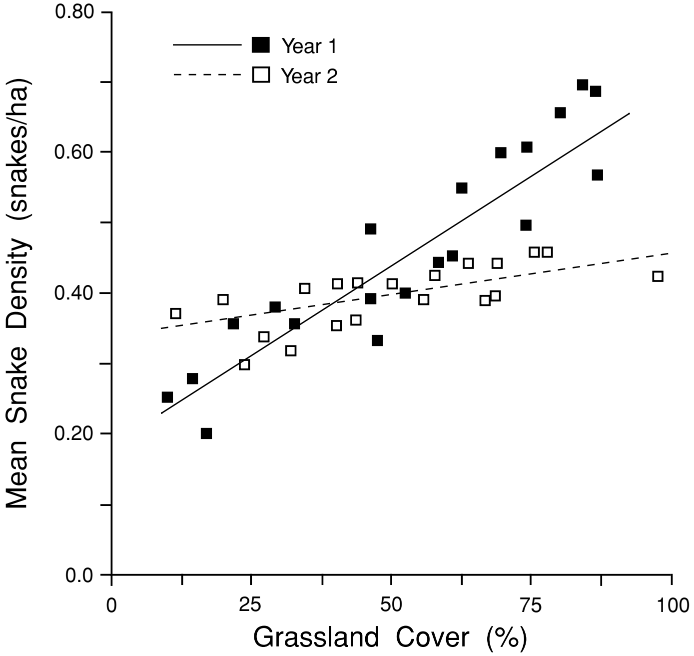


FIGURE 6.1. Relation between mean snake density and percent grassland cover in 40 landscapes based on hypothetical data. The non-parallel regression lines indicate a grassland cover × year interaction effect on mean snake density, which implies that the relation between mean snake density and grassland cover varied with year.

## 6.4. Assessing Temporal Variation in Animal–Landscape Relations Using Dependent Observations

### *6.4.1. Repeated-Measures Data in Landscape Studies*

Landscape ecologists frequently collect temporally repeated measures data. Typical situations include successive locations of radio-tagged animals, multiple observations of an organism's behavior, abundance or richness data collected at the same sites during successive years, and land-use change within a region. One rationale for collecting data on the same sampling units over time is that animal–landscape relations may vary temporally. By obtaining serial measurements from the same sampling units, one can develop an understanding of the nature and degree of this variation and incorporate it into uncertainty analyses. Another reason repeated measures data are collected is that sequential observations from the same units tend to have less variation than an equal number of observations from different sampling units, because intrinsic and extrinsic sources of variation are reduced. This condition can improve a landscape ecologist's ability to develop better inferences about the issue at hand because there is less noise that may obscure relations.

### *6.4.2. Statistical Nature of Repeated-Measures Data*

Data collected over time from the same sampling unit (organism, plot, landscape element, watershed, physiographic region) tend to be correlated. This violates the independent errors assumption of many statistical procedures for comparing the means of two or more groups (e.g., the *t*-test, standard analysis of variance) and for assessing relations between response variables and landscape characteristics (e.g., standard correlation and regression). Violations of this assumption can lead to artificially low standard errors, inflated Type I error rates, and hence spurious conclusions.

Common methods for analyzing repeated measures data are to conduct separate analyses for each time period, or to average responses across time periods. But these approaches avoid the temporal component entirely (Littell et al., 1998), do not permit simultaneous inference about both spatial and temporal components, and can result in less power because the sample size for one period of a study will be smaller than the sample size for all periods combined. Another option is to use a procedure that accommodates temporally correlated observations, such as traditional repeated-measures analysis of variance. However, this method requires that all pairs of measurements on a sampling unit are equally correlated regardless of the amount of elapsed time between observations (Littell et al., 1998), and that sets of observations taken at various points in time have equal variances. These conditions are rarely met in studies of animal–landscape relations. Observations on the same sampling unit taken close together in time are often more highly correlated than are observations obtained farther apart in time (Littell et al., 1998), and the variance of animal response variables often differs among time periods.

### *6.4.3. Advantages of Using Mixed Models to Analyze Repeated-Measures Data*

Development of general mixed models (Laird and Ware, 1982) has provided straightforward and flexible methods for assessing temporal dynamics of animal–landscape relations. Mixed models permit tests of fixed effects through either maximum likelihood (ML) or restricted maximum likelihood (REML) estimation. Temporal autocorrelation is accounted for by including temporal variables. The syntax for mixed models is similar to that of classic analysis of variance, and one can easily describe models, include interactions, and write code with basic SAS familiarity.

Mixed models represent a significant improvement over traditional repeated-measures analysis of variance in several ways:

- Mixed models allow for simultaneous inference about both spatial and temporal factors through the use of fixed and random effects.

- Mixed models apply more generally to a variety of covariance (correlation) structures and permit investigators to choose an appropriate covariance structure for the data being analyzed.
- Traditional repeated-measures analysis of variance does not readily allow for missing data. For example, if an observation for one individual is missing for one of the time periods, the data for all time periods for that individual must be excluded from the analysis, unless an estimate for the missing datum can be generated. Sometimes it is reasonable to do this by computing a mean based on the other observations in the same treatment group and time period, but this approach reduces the variance of the group and may thereby alter the outcome of the analysis in ways that are not defensible. Mixed models, on the other hand, accommodate incomplete records without the need for such estimates (Littell et al., 1998).

Landscape ecologists may include temporal effects in a mixed model for at least three reasons. One might be to control for effects of temporal variation. Adjusting parameter estimates, standard errors, and test statistics for temporal effects can prevent spurious conclusions and strengthen inferences. A second reason might be to examine potential interactions between time and spatial components. Although the effects of experimental treatments, landscape structure, or both are usually the primary concerns, understanding how these factors vary across time is often of interest as well. A third reason might be to identify the pattern of temporal correlation that best describes the data. For example, one might be interested in whether within-site correlations remain constant over time (compound symmetry) or whether these correlations decrease with time (autoregressive).

Below we demonstrate the basic approach for modeling temporally repeated-measures data with mixed models. We then demonstrate how to model more complex situations involving temporally and spatially dependent observations that landscape ecologists may encounter in analyses of temporal dynamics of animal–landscape relations. For more detailed instruction about mixed models than we provide here, we refer readers to guides for mixed models using SAS (Littell et al., 1996) or S-plus (Pinheiro and Bates, 2000).

### *6.4.4. Temporally Dependent Observations, Equally Spaced in Time*

When the same sampling units (landscapes, sites, individuals) are sampled over time at regular intervals (year, breeding season, week, day, etc.), the observations are equally spaced in time and are likely to be temporally dependent. Regular long-term monitoring of the same sites is a common source of such data. Sequential locations of radio-tagged individuals also may be temporally dependent; indeed, major radio-tracking texts (e.g., White and Garrott, 1990; Millspaugh and Marzluff, 2001) include discussion about the time interval between locations and independence of observations.

In many situations, mixed models can be used to model correlations between successive animal locations (e.g., Bowne et al., 1999). Mixed models can enable analysts to use more of the location data—because no observations have to be discarded—and to gain insight about the time interval within which successive locations are correlated. The example we describe next involves equally spaced repeated measures data and serves as a vehicle for describing the basic steps in analyzing mixed models.

### Example: Pine Siskin (*Carduelis pinus*) in Subalpine Forests

During each year of a 5-year experiment to assess effects of human intrusion on wildlife, birds were counted at 30 randomly placed permanent 1.0-ha sites in Wyoming subalpine forest (Fig. 6.2). Investigators randomly selected 20 of the 30 sites to receive experimental intrusions designed to mimic recreational disturbance by solitary hikers. The remaining 10 sites were unintruded controls. The treatments at these 30 permanent sites were the same during all 5 years, and $n$ for the entire study was $5 \times 30 = 150$. Full details of this experiment are available in Gutzwiller et al. (2002) and references therein.

For this example, we analyze the abundance of a small forest passerine, the pine siskin. This species' abundance at each site for a given year was calculated as the mean number of individuals detected during ten weekly point counts. Percentages of a 7.1-ha area around each site that were occupied by several land-cover types were estimated, but here we only use data for non-forested openings. Thus, for our example analysis, the important variables are pine siskin abundance, site, year, intrusion treatment, and percentage of the surrounding landscape covered by non-forested openings. Readers may obtain these data from the authors.

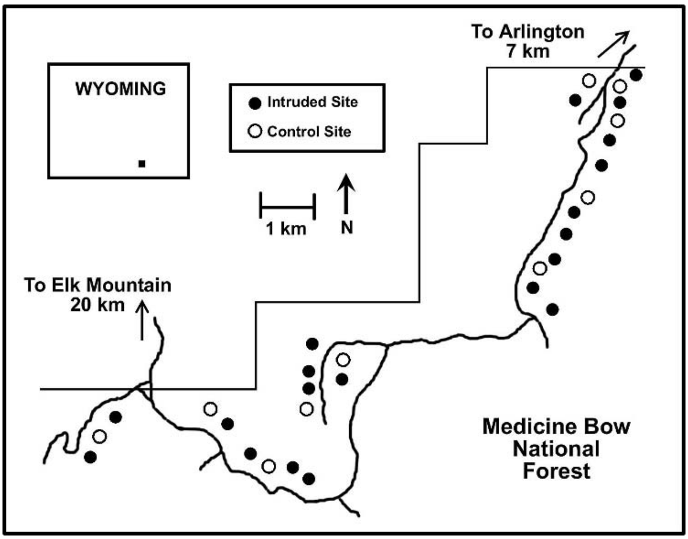


FIGURE 6.2. Map of study area in which pine siskin abundance was sampled during 1989–1993.

For purposes of demonstration, we asked two questions, one relating to a categorical landscape variable and another relating to a continuous landscape variable. Our first question was, "Does experimental intrusion, which mimicked human recreational activity in the landscape, increase or decrease pine siskin abundance?" Intrusion treatment was modeled as a categorical (1 or 0) variable. We consider a categorical variable in our example because evaluating animal–landscape relations often involves relating animal responses to categorical variables. Examples of such evaluations include comparing animal responses (e.g., nest success, movement rate, turning behavior, abundance) between edge sites vs. interior sites, between connected vs. unconnected patches, among different forest-cutting patterns, or among different landscape contexts.

Our second question, which involved a continuous variable, was, "Did pine siskin abundance vary with the percentage of the surrounding landscape covered by non-forested openings?" Landscape ecologists frequently investigate animal–landscape relations involving continuous landscape variables. Examples include species-area relations, relating animal metrics to gradients of urbanization or fragmentation, and using principal components of correlated landscape characteristics as predictor variables (e.g., Saab, 1999; Gutzwiller and Barrow, 2001; Riffell et al., 2003).

### Step 1: Specifying Fixed Effects

The first step in modeling repeated-measures data is to specify the fixed effects portion of the model. Temporal components are usually modeled using the `repeated` statement (see below). The initial model should contain all fixed main effects and interactions of interest. That is, the initial model should be a full model, and this full model should be fit using REML (Wolfinger, 1993). Using the pine siskin example, a program in SAS would look like this:

```
data pisi;
input site treat year pisi nfor;
cards;
1      0  1989  0.0  74
2      1  1989  0.0  64
. . .
. . .
29     1  1993  0.4  0
30     0  1993  0.5  0
run;
proc mixed method = reml;
class year treat site;
model pisi = treat nfor year treat*year nfor*year/
             ddfm = kenwardroger;
repeated year/subject = site(treat) type = cs r;
run;
```

This code enabled us to test whether pine siskin abundance (pisi) differed between intrusion treatments (treat), was related to amount of non-forested habitat in the surrounding landscape (nfor), and differed among years (year). To test whether the effects of intrusion and non-forested habitat varied from year to year, we used the interaction terms treat*year and nfor*year, respectively.

In the proc mixed statement above, the method=reml option requests restricted maximum likelihood estimation. In the model statement, the ddfm=kenwardroger option provides a small-sample adjustment of degrees of freedom for tests of fixed effects (Kenward and Roger, 1997). The repeated statement indicates that year is the repeated measure, and the options indicate the following specifications: subject=site(treat) specifies the individual sites (nested within treatment) as the subjects (sampling units) that are repeatedly sampled; type=cs specifies a compound symmetry covariance structure; and r causes printing of the estimated variance-covariance matrix (covariance matrix hereafter).

### Step 2: Selecting a Temporal Covariance Structure

Because we collected data on the same sites during each of 5 years, the five data points for a particular site may not have been independent. To address this issue, we can use a mixed model and consider the five yearly observations at each site to be repeated measures. Rather than ignoring or avoiding the implications of temporally correlated observations, the covariance structure of such data can be directly modeled, thereby supplying more detailed information about temporal dynamics of animal–landscape relations.

The ability to compare and select a covariance structure is a key advantage of the mixed model approach for repeated measures. SAS Proc Mixed offers over three dozen covariance structures (SAS Institute, 2002; see Table 6.1 in this chapter for five common examples), and they provide extraordinary flexibility in modeling temporal correlations. With so many possible structures, however, there is the possibility that selecting a covariance structure could become a "fishing expedition." To prevent this, we provide a general procedure for selecting an appropriate covariance structure (Box 6.2).

Selection should be done with two considerations in mind. First, are there any ecologically plausible temporal covariance structures? In our example below, we considered the possibility that abundances from the same site were equally correlated (perhaps because of habitat or environmental similarities) regardless of the number of years between pairs of observations; this structure is referred to as compound symmetry. We also considered an autoregressive covariance structure, which represented the possibility that repeated measures of pine siskin abundance obtained closer together in time would be more highly correlated than would observations made farther apart in time. Many passerine species are faithful to breeding sites from one year to the next, but these species typically live only a few years. Thus, abundance estimates from two successive years may involve some of the same individual birds, but estimates obtained more than 1 year apart may involve

**Box 6.2. General procedure for selecting a covariance structure.**

1. Fit the fixed effects portion of the model.
2. Identify a set of candidate covariance structures.
   - Consider ecological and biological characteristics of the dependent variables. For example, consider whether the biology of the organism suggests that the variance of the response variable might fluctuate from year to year.
   - Consider parsimony of the covariance structure relative to available sample size. Many of the available covariance structures require a large number of extra parameters, which may exceed the number of parameters that can be confidently estimated for a given sample size.
3. Fit a *separate* mixed model (with an identical fixed effects portion) using each of the candidate covariance structures.
4. Select the most appropriate covariance structure using one or more model fit statistics such as Akaike's Information Criterion (AIC) or Schwarz's Bayesian Criterion (SBC or BIC).

increasingly higher proportions of new individuals, reducing potential correlation in responses.

Another characteristic of many animal populations is the potential for both the abundance and the variance in abundance to fluctuate from year to year. SAS includes heterogeneous versions (Wolfinger, 1996) of the compound symmetry and autoregressive structures. Heterogeneous structures may be useful for modeling variable populations because they allow the diagonals of the covariance structure (the yearly variances in our pine siskin example) to be different each year (see Table 6.1), unlike the standard compound symmetry, autoregressive, and many other structures. We therefore evaluated heterogeneous versions of these covariance structures in the analysis presented below.

A second consideration for choosing a temporal covariance structure should be the number of additional parameters that a particular structure will require. In our pine siskin example, the unstructured covariance structure would have required the estimation of 15 parameters (Table 6.1) just for the temporal covariance part of the model. Including fixed effects and the intercept, we would have had considerably fewer than 10 observations per parameter (our total $n$ was 150). Having few observations relative to the number of parameters can decrease power and increase the probability of spurious effects (Flack and Chang, 1987; Morrison et al., 1998; Burnham and Anderson, 2002). Accordingly, we did not consider the unstructured covariance structure or other covariance structures with a large number of parameters.

This left us with a set of four candidate covariance structures (number of parameters in parentheses): compound symmetry (2), heterogeneous compound symmetry ($t + 1 = 6$, where $t$ = number of time intervals), autoregressive (2), and heterogeneous autoregressive ($t + 1 = 6$). To evaluate these four covariance structures,

TABLE 6.1 Examples of five covariance structures available in SAS Proc Mixed.

| | |
|---|---|
| Compound Symmetry (CS) | $\begin{bmatrix} \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 \end{bmatrix}$ |
| Heterogeneous Compound Symmetry (CSH) | $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho & \sigma_1\sigma_5\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho & \sigma_2\sigma_5\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 & \sigma_4\sigma_5\rho \\ \sigma_5\sigma_1\rho & \sigma_5\sigma_2\rho & \sigma_5\sigma_3\rho & \sigma_5\sigma_4\rho & \sigma_5^2 \end{bmatrix}$ |
| Autoregressive (AR[1]) | $\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$ |
| Heterogeneous Autoregressive (ARH[1]) | $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 & \sigma_1\sigma_5\rho^4 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 & \sigma_2\sigma_5\rho^3 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho^2 \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 & \sigma_4\sigma_5\rho \\ \sigma_5\sigma_1\rho^4 & \sigma_5\sigma_2\rho^3 & \sigma_5\sigma_3\rho^2 & \sigma_5\sigma_4\rho & \sigma_5^2 \end{bmatrix}$ |
| Unstructured (UN) | $\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_5^2 \end{bmatrix}$ |

we used four separate runs of Proc Mixed. Each time we fit the same fixed effects portion of the model but modified the `repeated` statement to include a different covariance structure. We show the different `repeated` statements below; the rest of the SAS code remained exactly as shown previously.

```
repeated year/subject=site(treat) type=cs r;
repeated year/subject=site(treat) type=csh r;
repeated year/subject=site(treat) type=ar(1) r;
repeated year/subject=site(treat) type=arh(1) r;
```

SAS provides several model fit statistics that can be used to select a covariance structure. Of these, Akaike's Information Criterion (AIC) and Schwarz's Bayesian Criterion (BIC or sometimes SBC) are common choices (Littell et al., 1996; Wolfinger, 1996, 1997). Both are based on the log likelihood and include a penalty proportional to the number of covariance parameters (BIC provides a stiffer penalty than does AIC). Burnham and Anderson (2002) recommend using the small-sample version of Akaike's Information Criterion ($AIC_c$, also provided by SAS) when $n < 40$ per estimated parameter in a model (including all covariance parameters, fixed effect parameters, intercept, and error terms). We used $AIC_c$ to select the most appropriate covariance structure for our example.

$AIC_c$ (smaller is better) indicated that the heterogeneous compound symmetry covariance structure was the most appropriate choice, but that the heterogeneous autoregressive structure was a very close second (difference in $AIC_c < 0.1$).

| Covariance structure | $AIC_c$ |
|---|---|
| Compound symmetry | 22.7 |
| Heterogeneous compound symmetry | 16.4 |
| Autoregressive | 23.3 |
| Heterogeneous autoregressive | 16.5 |

When two or more covariance structures have similar measures of appropriateness, it is not clear which structure is more appropriate. If the primary goal is to improve inference about fixed effects, the choice among appropriate covariance structures does not present a major dilemma. Littell et al. (1996, p. 321) note that, "the major impact on inference results from using a *reasonable* covariance model. The specific model used is not nearly as important, as long as it is 'in the ballpark.'"

In many studies, experimental treatments may affect the variance instead of (or in addition to) the mean. One option in this situation is to used the `group=` option in the `repeated` statement, which permits different values for each parameter in the covariance structure for each level of the `group` effect (SAS Institute, 2002). The repeated statement would look like the following:

```
repeated year/subject = site(treat) type = csh group = treat r;
```

In our current pine siskin example, both intrusion treatments would exhibit heterogeneous compound symmetry structure but the variances (diagonals) could be different for sites in each treatment. The `group=` option should be used with caution, especially with small sample sizes, because it greatly increases the number of parameters in and the complexity of the covariance structure.

## Step 3: Inference about Dynamics of Animal–Landscape Relations

The third step involves making the statistical inference using the previously selected covariance structure. Thus, we used the following SAS statements. Note that the `type =` option in the `repeated` statement is set to `csh` for heterogeneous compound symmetry.

```
data pisi;
input site treat year pisi nfor;
cards;
1      0  1989  0.0  74
2      1  1989  0.0  64
. . .
. . .
29     1  1993  0.4  0
30     0  1993  0.5  0
run;
proc mixed method = reml;
class year treat site;
model pisi = treat nfor year treat*year nfor*year/
             ddfm = kenwardroger;
repeated year/subject = site(treat) type = csh r;
run;
```

One useful output from this program is the table of parameters for the covariance matrix:

Covariance parameter estimates

| Covariance parameter | Subject | Estimate |
|---|---|---|
| Var(1) | site(treat) | 0.0173 |
| Var(2) | site(treat) | 0.0264 |
| Var(3) | site(treat) | 0.0389 |
| Var(4) | site(treat) | 0.0682 |
| Var(5) | site(treat) | 0.0497 |
| CSH | site(treat) | 0.1205 |

In this output in the Estimate column, the yearly variances in pine siskin abundance (i.e., Var(1) for 1989 = 0.0173, Var(2) for 1990 = 0.0264, etc.) are the diagonals of the CSH covariance matrix (see Table 6.1 and Box 6.3). The variances generally increase with year, confirming the choice of heterogeneous structures. The CSH parameter in the Estimate column is the constant ($\rho$) in the covariance part of the matrix (see Table 6.1 and Box 6.3). Overall, these results imply that heterogeneous compound symmetry was a reasonable covariance structure for our data.

Next, we look at the tests of fixed effects.

Type 3 tests of fixed effects

| Effect | Numerator DF | Denominator DF | $F$ | Pr > $F$ |
|---|---|---|---|---|
| treat | 1 | 26.9 | 0.10 | 0.7582 |
| year | 4 | 58.2 | 15.75 | <0.0001 |
| nfor | 1 | 26.9 | 4.24 | 0.0493 |
| year*treat | 4 | 58.2 | 0.95 | 0.4424 |
| nfor*year | 4 | 58.2 | 3.42 | 0.0140 |

**Box 6.3. Example calculations for a covariance structure.**

Generic heterogeneous compound symmetry (CSH) structure (in the first matrix below) compared to the specific CSH structure parameterized for the pine siskin abundance data. Example calculations for the first column of covariances are presented in the second matrix. The specific structure (third matrix) was generated by the analysis described in 6.4.4. *Temporally Dependent Observations, Equally Spaced in Time* (Step 3: Inference about Dynamics of Animal–Landscape Relations).

Generic CSH Structure

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho & \sigma_1\sigma_5\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho & \sigma_2\sigma_5\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 & \sigma_4\sigma_5\rho \\ \sigma_5\sigma_1\rho & \sigma_5\sigma_2\rho & \sigma_5\sigma_3\rho & \sigma_5\sigma_4\rho & \sigma_5^2 \end{bmatrix}$$

Example Calculations for Specific CSH Structure (see text)

$$\begin{bmatrix} 0.0173 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho & \sigma_1\sigma_5\rho \\ \sqrt{0.0264}\sqrt{0.0173}\ (0.1205) & 0.0264 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho & \sigma_2\sigma_5\rho \\ \sqrt{0.0389}\sqrt{0.0173}\ (0.1205) & \sigma_3\sigma_2\rho & 0.0389 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho \\ \sqrt{0.0682}\sqrt{0.0173}\ (0.1205) & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & 0.0682 & \sigma_4\sigma_5\rho \\ \sqrt{0.0497}\sqrt{0.0173}\ (0.1205) & \sigma_5\sigma_2\rho & \sigma_5\sigma_3\rho & \sigma_5\sigma_4\rho & 0.0497 \end{bmatrix}$$

Specific CSH Structure†

$$\begin{bmatrix} 0.0173 & 0.002573 & 0.003127 & 0.004139 & 0.003533 \\ 0.002573 & 0.0264 & 0.003858 & 0.005108 & 0.004360 \\ 0.003127 & 0.003858 & 0.0389 & 0.006207 & 0.005298 \\ 0.004139 & 0.005108 & 0.006207 & 0.0682 & 0.007013 \\ 0.003533 & 0.004360 & 0.005298 & 0.007013 & 0.0497 \end{bmatrix}$$

† To produce the specific covariance structure as SAS output, use the `r` option in the `repeated` statement of Proc Mixed.

In any multifactor model with interactions, one should first check for significant interactions (Wolfinger, 1997) because when an interaction effect is present, the influence of one main effect depends on the level of the other main effect involved in the interaction (Underwood, 1997). The `nfor*year` interaction was significant (see above), so we focused our interpretation on the `nfor*year` interaction instead of on the associated main effects. Our conclusions, based on the table above, were: pine siskin abundance was not lower on intruded sites; pine siskin abundance differed significantly among years, but this effect varied with the percent of the surrounding landscape in non-forested area; and pine siskin abundance was related to the percent of non-forested area in the surrounding landscape, but this relation varied among years. These latter two interpretations are alternate ways of viewing the `nfor*year` interaction effect.

To explore the temporal dynamics of the relation between `pisi` and `nfor` further, we can modify our mixed-model code so it produces intercepts and slopes for the relation between pine siskin abundance and non-forested openings for each year separately. This is accomplished by dropping the `nfor` main effect and other non-significant terms, and by specifying two options in the `model` statement: `noint` tells Proc Mixed not to fit a common intercept, but five separate intercepts (one for each year); `solution` requests the estimates for all of the fixed-effect parameters.

```
data pisi;
input site treat year pisi nfor;
cards;
1      0  1989  0.0  74
2      1  1989  0.0  64
. . .
. . .
29     1  1993  0.4  0
30     0  1993  0.5  0
run;
proc mixed method=reml;
class year treat site;
model pisi-year nfor*year/ddfm=kenwardroger noint
             solution;
repeated year/subject=site(treat) type=csh r;
run;
```

This program generated the following output:

Solutions for fixed effects

| Effect | Date | Estimate | Standard error | DF | *t* | Pr > \|*t*\| |
|---|---|---|---|---|---|---|
| year | 1989 | 0.0731 | 0.0287 | 28.2 | 2.55 | 0.0166 |
| year | 1990 | 0.0901 | 0.0342 | 29.2 | 2.63 | 0.0134 |
| year | 1991 | 0.1593 | 0.0415 | 28.9 | 3.84 | 0.0006 |
| year | 1992 | 0.4157 | 0.0560 | 27.3 | 7.43 | <0.0001 |
| year | 1993 | 0.4326 | 0.0479 | 27.0 | 9.03 | <0.0001 |
| nfor*year | 1989 | −0.0009 | 0.0013 | 28.2 | −0.65 | 0.5222 |
| nfor*year | 1990 | 0.0052 | 0.0016 | 29.2 | 3.35 | 0.0022 |
| nfor*year | 1991 | 0.0050 | 0.0019 | 28.9 | 2.65 | 0.0130 |
| nfor*year | 1992 | 0.0007 | 0.0026 | 27.3 | 0.26 | 0.7964 |
| nfor*year | 1993 | 0.0009 | 0.0022 | 27.0 | 0.43 | 0.6725 |

In the Estimate column above, the coefficients for `year` are intercepts, and the coefficients for `nfor*year` are slopes. The intercept estimates indicate that pine siskin abundance and its standard error generally increased over time. The slope estimates indicate the relation between pine siskin abundance and nfor was strongly positive in 1990 and 1991, as indicated by the significant *t* statistics, but not during the other 3 years. Using these results, one can explore, through theoretical arguments or further experiments, why this relation was significant in these 2 years but not the others.

### *6.4.5. Temporally Dependent Observations, Unequally Spaced in Time*

Temporally repeated observations often occur at irregular intervals. Consider the following hypothetical example of an unequally spaced design: an investigative team planted vegetative corridors in different matrix types and then monitored animal use of those corridors 1 year, 2 years, 4 years, 7 years, and 10 years later. A situation like this could arise simply because of funding or other logistical constraints that prohibit sampling at equal intervals. Unequally spaced observations also may occur when weather conditions restrict sampling to specific but unpredictable times, or when a marked animal is relocated after a period of being undetectable.

With unequally spaced temporal observations, landscape ecologists can still use mixed models, but temporal correlations must be modeled as a function of "distance" rather than as a function of a regular time interval. In this situation, "distance" is the single dimension of time, not two- or three-dimensional space, and structures commonly used to model spatial covariance (power law, Gaussian, spherical, etc.) are used to model temporal covariance (Littell et al., 1996). Landscape ecologists have fully recognized the utility of spatial covariance structures for modeling spatial variation in animal–landscape relations (e.g., Selmi and Boulinier, 2001; Keitt et al., 2002; Evans and Gaston, 2005), but the same covariance structures, and very similar SAS code, also can be used to model covariance among unequally spaced, temporally correlated observations. Recent versions of SAS include over a dozen spatial covariance parameters that can be included in the `type=` option of the `repeated` command in Proc Mixed.

Although our pine siskin data contain equally spaced repeated measures, we used it as an example of how to code for unequally spaced repeated measures below. We used a power law spatial covariance, which provides a generalization of the autoregressive (AR[1]) structure for equally spaced data.

```
data pisi;
input site treat year pisi nfor;
cards;
1      0   1989  0.0  74
2      1   1989  0.0  64
. . .
. . .
29     1   1993  0.4  0
30     0   1993  0.5  0
run;
data pisi2; set pisi;
year1=year;
run;
proc mixed method=reml data=pisi2 order=data;
class year treat site;
```

```
model pisi = treat nfor year treat*year nfor*year/
              ddfm=kenwardroger;
repeated year/subject=site(treat)
  type=sp(pow)(year1) r;
run;
```

An important caveat is that the spatial covariance analysis requires that year be a *continuous* variable in the `type=` option of the `repeated` statement. Using a `data` step, we created a second time variable, `year1`, that was identical to the categorical `year` variable, except that it was considered continuous (notice that `year1` is not in the `class` statement).

### *6.4.6. Temporally and Spatially Dependent Observations, Two Dimensions*

Research efforts on animal–landscape relations often result in data that are correlated through both space and time. For example, if invertebrates were sampled in multiple stream reaches in each of several watersheds over time, there might be correlations among reaches located in the same watersheds (spatial dependence) and correlations among temporal observations in the same reaches (temporal dependence). Another example would involve serial sampling of the same patches over time in distinct physiographic regions. A third example would be repeated location data on individual animals that form groups in different areas.

#### Pine Siskin Example: Doubly Repeated Measures

Returning to the pine siskin example, Fig. 6.2 indicates that the 30 sites occurred in two basic groups, one in the southwestern part of the study area and one in the northeastern part of the study area. These two groups corresponded to two areas that were relatively snow-free and hence accessible during the early part of the breeding season. The sites were positioned randomly, treatments were randomly assigned to sites, and there were no major vegetation differences between treatment groups. For demonstration purposes, we assume that in addition to the temporally repeated measures associated with year, the two groups (or clusters) of sites involve spatially repeated measures within each cluster. Thus, our challenge now is to simultaneously model the correlation among temporally repeated measures and the correlation among sites within a cluster.

Doubly repeated measures can be dealt with in mixed models by using the `repeated` statement, the `random` statement, or both. In addition to the temporal effect `year`, which is specified by the `repeated` statement, we can assign group membership to a categorical variable (coded 1 or 2), named `cluster` here, and model the spatial dependence as a random effect:

```
data pisi;
input site treat year pisi nfor cluster;
```

```
cards;
1      0  1989  0.0  74  1
2      1  1989  0.0  64  1
. . .
. . .
29     1  1993  0.4  0   2
30     0  1993  0.5  0   2
run;
class year treat site cluster;
model pisi=treat nfor year treat*year nfor*year/
model pisi = treat nfor year treat*year nfor*year/
              ddfm=kenwardroger;
random cluster;
repeated year/subject=site(treat) type=csh r;
run;
```

This program generated the following output.

Covariance parameter estimates

| Covariance parameter | Subject | Estimate |
|---|---|---|
| Cluster | — | 0.0072 |
| Var(1) | site(treat) | 0.0200 |
| Var(2) | site(treat) | 0.0184 |
| Var(3) | site(treat) | 0.0284 |
| Var(4) | site(treat) | 0.0652 |
| Var(5) | site(treat) | 0.0599 |
| CSH | site(treat) | 0.0799 |

In this output, as before, Var(1) through Var(5) in the Estimate column are the yearly variances in pine siskin abundance (diagonals of the covariance matrix), and the CSH parameter is the constant in the covariance part of the matrix (Table 6.1). Notice in the Estimate column for Cluster that the covariance associated with the groups of sites was an order of magnitude smaller than were the other covariance parameters. Furthermore, the fixed-effect results (not shown) did not change appreciably, so in this example the spatial grouping of the sites was not important.

### *6.4.7. Temporally and Spatially Dependent Observations, Three Dimensions*

Three-dimensional repeated-measures data—a temporal correlation (repeated measures over time) and correlation in two spatial dimensions (typically x and y geographic coordinates)—often arise when studying animal–landscape relations. This situation occurs when there is a set of permanent sampling stations located throughout a landscape or region, and these stations are sampled repeatedly over time. Animal metrics measured at nearby stations may be more correlated than are those for distant stations, and animal metrics measured close together in time may be more correlated than are those measured farther apart in time. Furthermore, spatial correlations may change over time, and temporal correlations may change through space. Examples of three-dimensional repeated-measures data include the 30 stations at which investigators repeatedly sampled pine siskins, and the thousands of Breeding Bird Survey routes sampled annually across North America (Robbins et al., 1986; Sauer et al., 2005).

Although Proc Mixed would allow a temporal variable to be treated as a third spatial dimension (e.g., `type=sp(pow)(easting northing year1)`), this is not appropriate because time and space do not have comparable units (Schabenberger and Gotway, 2005). Other possible approaches for three-dimensional situations would be to either conduct separate spatial analyses for each level of the time dimension, or conduct separate temporal analyses for each location. However, these approaches do not account for possible interactions between spatial and temporal processes. The ideal approach would be to model the spatial and temporal correlations and space–time interactions simultaneously, but techniques for doing this are not well-developed or readily accessible in common statistical packages.

If certain assumptions are met, one can analyze 3-dimensional repeated-measures data using either separable covariance structures, or non-separable covariance structures. Separable covariance structures permit joint analysis of spatio-temporal data, but do not permit space-time interactions to be investigated (Mitchell and Gumpertz, 2003; Schabenberger and Gotway, 2005), whereas techniques based on non-separable covariance structures allow for both joint and interaction analyses (Schabenberger and Gotway, 2005). Unfortunately, use of these two types of structures is complex, and writing code for the analyses is not simple in popular statistical packages (but see Mitchell and Gumpertz, 2003 for a spatio-temporal analysis using several SAS procedures). The references cited above are an excellent starting point for researchers interested in pursuing these techniques.

### *6.4.8. Summary of Mixed Models for Repeated Measures*

Mixed models provide an opportunity to explicitly incorporate simultaneous inference about time and space in studies of animal–landscape relations. These models are flexible, allowing temporal and spatial effects to be addressed in different ways and with different levels of dimensionality.

For our inferences about pine siskin, the mixed-model approach was beneficial in three major ways. First, it enabled us to estimate effects of human intrusion treatments and percent of the surrounding landscape composed of non-forested openings after accounting for different covariance structures in the data, which reduced the possibility of spurious conclusions. Second, it provided the ability to use all of the observations in a single analysis and thereby avoid the loss of statistical power that might have been incurred by splitting the dataset and conducting analyses for each year separately. Finally, the mixed-model approach enabled us to make simultaneous inferences about spatial and temporal factors.

The basic protocol for conducting mixed-model analyses in the context of temporally repeated measures can be summarized as follows:

Step 1: Fit the fixed effects portion of the model using REML. Generally, this part of the model should contain all main effects and interactions of interest.

Step 2: With thought to ecological processes and sample size restrictions, choose a set of candidate covariance structures and select the best-fitting structure using a model fit statistic such as AIC.

Step 3: Apply the selected covariance structure for inference in the final model using REML.

Note that after selecting the appropriate covariance structure, investigators may wish to evaluate several different competing fixed-effect models. Often, AIC (or another model selection criterion) is used to select the "best" model(s) for inference (Burnham and Anderson, 2002). There are two important considerations when doing this. First, investigators should use maximum likelihood (use the `method=ml` option in the `proc mixed` statement) when comparing models with different fixed-effect specifications, because REML restricts the optimization of the likelihood function to the random-effects portion of the model (Wolfinger, 1993). Second, one should verify that SAS Proc Mixed uses the number of parameters (including intercepts and error terms) specified in Burnham and Anderson (2002) for AIC calculations (see Stafford and Strickland, 2003).

## *6.4.9. Additional Information About Mixed Models*

### Mixed-Model Diagnostics

For classical linear modeling approaches like regression and analysis of variance, various tools are available to assess the overall fit of the model to the data and to examine the influence of individual observations on the model. These tools include residual analysis, collinearity analysis, and influence analysis. Such analyses can be useful for assessing the degree to which model assumptions are met and identifying individual data points that have a strong influence on structuring the model. In mixed models for repeated measures, these assessments are more difficult because it is not the influence of individual observations (e.g., a particular pine siskin survey) that is of interest, but rather the influence of a particular site that was observed multiple times (e.g., over several years). Recent versions of SAS include options to produce influence diagnostics that allow assessment of the fit of both random- and fixed-effect components (Schabenberger, 2004).

### Mixed Models for Non-Traditional Data

We have focused on linear models involving continuous response variables that are normally distributed, but the mixed-modeling concepts and techniques we have described can be extended to other types of response variables. Mixed models can be applied to binomial or Poisson distributions via Proc Glimmix (Littell et al., 1996; Schabenberger, 2005) using syntax and theory that are similar to those described above. Mixed-model approaches can be extended to non-linear models via Proc Nlinmixed in SAS (Littell et al., 1996), or with S-Plus (Pinheiro and Bates, 2000). Detailed discussion of these options is beyond the scope of this chapter, but we mention them for readers with interests in applications to binary and count data and to non-linear processes. The references cited above are good portals into the pertinent literature.

## 6.5. Conclusions

Several established modeling methods accessible to those with moderate training in statistics can be used to incorporate time into studies of animal–landscape relations. Typical advantages of explicitly modeling time in such relations include: results that are more defensible on technical grounds; better understanding of the ecology involved; knowledge of the magnitude of temporal variation in the relations, which can be used to characterize temporal flux in, and level of uncertainty about, the relations; and robust predictions about animal use of landscapes over time. Knowledge about temporal variation in animal–landscape relations also can be used to parameterize and structure simulation models (Gutzwiller and Barrow, 2001).

These advantages hold promise for advancing the disciplines of landscape ecology, land-use planning, and biological conservation for the following reasons. Defensible results are crucial for establishing policy and management guidelines. The value of a model lies largely in the ecological understanding it provides, and models that address temporal flux are more likely to provide better ecological understanding than are models that that do not account for such dynamics. The utility of an animal–landscape model can be constrained by uncertainty about whether it holds over time, and explicitly modeling temporal dynamics can help one identify levels of uncertainty. Predictions from models that consider temporal dimensions also are likely to be more robust because they probably represent reality more accurately. Simulation models are frequently used to predict consequences of environmental disturbances and management decisions. By incorporating temporally explicit statistical models as key components (sub-models), simulation models may represent temporal dynamics more realistically.

*Acknowledgments*. We thank J. A. Bissonette, M. W. Schilling, I. Storch, and B. Strickland for advice about the manuscript.

## *References*

Bowne, D.R., Peles, J.D., and Barrett, G.W. 1999. Effects of landscape spatial structure on movement patterns of the hispid cotton rat (*Sigmodon hispidus*). Landscape Ecology 14:53–65.

Burnham, K.P., and Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second Edition. New York: Springer-Verlag.

Evans, K.L., and Gaston, K.J. 2005. People, energy, and avian species richness. Global Ecology and Biogeography 14:187–196.

Flack, V.F., and Chang, P.C. 1987. Frequency of selecting noise variables in subset regression analysis: A simulation study. American Statistician 41:84–86.

Gutzwiller, K.J., ed. 2002. Applying Landscape Ecology in Biological Conservation. New York: Springer-Verlag.

Gutzwiller, K.J., and Barrow, W.C. Jr. 2001. Bird–landscape relations in the Chihuahuan Desert: Coping with uncertainties about predictive models. Ecological Applications 11:1517–1532.

Gutzwiller, K.J., and Barrow, W.C. Jr. 2002. Does bird community structure vary with landscape patchiness? A Chihuahuan Desert perspective. Oikos 98:284–298.

Gutzwiller, K.J., Riffell, S.K., and Anderson, S.H. 2002. Repeated human intrusion and the potential for nest predation by gray jays. Journal of Wildlife Management 66:372–380.

Keitt, T.H., Bjørnstad, O.N., Dixon, P.M., and Citron-Pousty, S. 2002. Accounting for spatial pattern when modeling organism-environment interactions. Ecography 25:616–625.

Kenward, M.G., and Roger, J.H. 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997.

Laird, N.M., and Ware, J.H. 1982. Random-effects models of longitudinal data. Biometrics 38:963–974.

Littell, R.C., Henry, P.R., and Ammerman, C.B. 1998. Statistical analysis of repeated measures data using SAS procedures. Journal of Animal Science 76:1216–1231.

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. 1996. SAS System for Mixed Models. Cary, NC: SAS Institute.

McCoy, T.D., Ryan, M.R., Kurzejeski, E.W., and Burger, L.W. Jr. 1999. Conservation Reserve Program: Source or sink habitat for grassland birds in Missouri? Journal of Wildlife Management 63:530–538.

McGarigal, K., and McComb, W.C. 1995. Relationships between landscape structure and breeding birds in the Oregon Coast Range. Ecological Monographs 65:235–260.

Millspaugh, J.J., and Marzluff, J.M., eds. 2001. Radio Tracking and Animal Populations. San Diego: Academic Press.

Mitchell, M.W., and Gumpertz, M.L. 2003. Spatio-temporal prediction inside a free-air $CO_2$ enrichment system. Journal of Agricultural, Biological, and Environmental Statistics. 8:310–327.

Morrison, M.L. 2002. Role of temporal and spatial scale. Pages 123–124 *in* Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., and Samson, F.B., eds. Predicting Species Occurrences: Issues of Accuracy and Scale. Washington, DC: Island Press.

Morrison, M.L., Marcot, B.G., and Mannan, R.W. 1998. Wildlife-Habitat Relationships: Concepts and Applications. Madison: University of Wisconsin Press.

Naugle, D.E., Higgins, K.F., Nusser, S.M., and Johnson, W.C. 1999. Scale-dependent habitat use in three species of prairie wetland birds. Landscape Ecology 14:267–276.

Neter, J., Wasserman, W., and Kutner, M.H. 1989. Applied Linear Regression Models. Second Edition. Burr Ridge, IL: Richard D. Irwin.

O'Connor, R.J. 1986. Dynamical aspects of avian habitat use. Pages 235–240 *in* Verner, J., Morrison, M.L., and Ralph, C.J., eds. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. Madison: University of Wisconsin Press.

Pearson, S.M. 1993. The spatial extent and relative influence of landscape-level factors on wintering bird populations. Landscape Ecology 8:3–18.

Pinheiro, J.C., and Bates, D.M. 2000. Mixed Effects Models in S and S-plus. New York: Springer-Verlag.

Riffell, S.K., Keas, B.E., and Burton, T.M. 2003. Birds in Great Lakes coastal wet meadows: Is landscape context important? Landscape Ecology 18:95–111.

Robbins, C.S., Bystrack, D., and Geissler, P.H. 1986. The Breeding Bird Survey: Its First Fifteen Years, 1965–1979. Resource Publication 157. Washington, DC: U.S. Fish and Wildlife Service.

Saab, V. 1999. Importance of spatial scale to habitat use by breeding birds in riparian forests: A hierarchical analysis. Ecological Applications 9:135–151.

SAS Institute. 2002. SAS OnlineDoc 9. Cary, NC: SAS Institute. http://v9doc.sas.com/sasdoc/ (last accessed 6 October 2006).

Sauer, J.R., Hines, J.E., and Fallon, J. 2005. The North American Breeding Bird Survey, Results and Analysis 1966–2004. Version 2005.2. Laurel, MD: USGS Patuxent Wildlife Research Center. http://www.mbr-pwrc.usgs.gov/bbs/ (last accessed 6 October 2006).

Schabenberger, O. 2004. Mixed model influence diagnostics. Proceedings of the 29th Annual SAS Users Group International Conference. Cary, NC: SAS Institute.

Schabenberger, O. 2005. Introducing the GLIMMIX procedure for generalized linear mixed models. Proceedings of the 30th Annual SAS Users Group International Conference. Cary, NC: SAS Institute.

Schabenberger, O., and Gotway, C.A. 2005. Statistical Methods for Spatial Data Analysis. Boca Raton, FL: Chapman & Hall/CRC Press.

Shochat, E., and Tsurim, I. 2004. Winter bird communities in the northern Negev: Species dispersal patterns, habitat use and implications for habitat conservation. Biodiversity and Conservation 13:1571–1590.

Selmi, S., and Boulinier, T. 2001. Ecological biogeography of southern ocean islands: The importance of considering spatial issues. American Naturalist 158:426–437.

Stafford, J.D., and Strickland, B.K. 2003. Potential inconsistencies when computing Akaike's information criterion. Bulletin of the Ecological Society of America 84:68–69.

Underwood, A.J. 1997. Experiments in Ecology. Cambridge, United Kingdom: Cambridge University Press.

White, G.C., and Garrott, R.A. 1990. Analysis of Wildlife Radio-Tracking Data. San Diego: Academic Press.

Wolfinger, R.D. 1993. Covariance structure selection in general mixed models. Communications in Statistics, Simulation and Computation 22:1079–1106.

Wolfinger, R.D. 1996. Heterogeneous variance-covariance structures for repeated measures. Journal of Agricultural, Biological, and Environmental Statistics 1:205–230.

Wolfinger, R.D. 1997. An example of using mixed models and Proc Mixed for longitudinal data. Journal of Biopharmaceutical Statistics 7:481–500.